

Revised Draft: September 2015

# Bodily Action and Distal Attribution in Sensory Substitution

Robert Briscoe

Department of Philosophy, Ohio University ([rbriscoe@gmail.com](mailto:rbriscoe@gmail.com))

**Abstract:** According to proponents of the sensorimotor contingency theory of perception (Hurley & Noë 2003, Noë 2004, O'Regan 2011), active control of camera movement is necessary for the emergence of distal attribution in tactile-visual sensory substitution (TVSS) because it enables the subject to acquire knowledge of the way stimulation in the substituting modality varies as a function of self-initiated, bodily action. This chapter, by contrast, approaches distal attribution as a solution to a *causal inference* problem faced by the subject's perceptual systems. Given all of the available endogenous and exogenous evidence available to those systems, what is the most probable source of stimulation in the substituting modality? From this perspective, active control over the camera's movements matters for rather different reasons. Most importantly, it generates proprioceptive and efference-copy based information about the camera's body-relative position necessary to make use of the spatial cues present in the stimulation that the subject receives for purposes of egocentric object localization.

**Keywords:** Bayesian perception, distal attribution, egocentric space, enactivism, neural plasticity, proprioception, sensory substitution, sensorimotor contingencies, spatial representation

## 1. Two Questions about Sensory Substitution

Sensory substitution devices (SSDs) convert images from a video camera into patterns of vibrotactile stimulation (White *et al.* 1970; Bach-y-Rita 1972, 2004), electrotactile stimulation (Sampaio *et al.* 2001, Püto *et al.* 2005, Chebat *et al.* 2011), or auditory stimulation (Meijer 1992, Capelle *et al.* 1998, Renier *et al.* 2005, Amedi *et al.* 2007, Kim & Zatorre 2008) that visually-impaired individuals can use to perform tasks ordinarily guided by non-prosthetic vision. An adequate account of how SSDs enable trained users to interact with the environment in an adaptive manner must answer two main questions. The first question concerns the end-product of the learning process in sensory substitution. What kind (or kinds) of representational states do properly trained subjects form in response to stimulation in the substituting perceptual modality? There are at least five possibilities:

- a) The SSD extends the range of properties represented by the *substituting* modality. The device, for example, enables trained users to perceive the shapes and sizes of objects at a distance from their body by means of touch.

- b) The SSD enables trained subjects to perceive properties of the distal environment via the *substituted* modality, i.e., the device to a certain extent restores the sense of sight.
- c) The SSD enables trained subjects to perceive properties of the distal environment via a *new* prosthetic modality, i.e., the subject enjoys experiences that are not straightforwardly in the substituting or substituted modality.
- d) Trained subjects transform stimulation in the substituting modality into accurate visual and/or spatial *mental images* of surrounding objects and their properties.
- e) Trained subjects perform quick *cognitive inferences* from stimulation in the substituting modality to properties of the distal environment.

The second question, by contrast, concerns the nature of the *learning process* itself. How do blindfolded, late blind, and early blind subjects respectively learn to use the stimulations provided by a given type of SSD for such purposes as object-directed motor control, wayfinding, or object recognition? It is on this second question that I shall focus primarily in this chapter.

Before proceeding, it should be emphasized that both questions need to be relativized not only to the specific type of SSD being deployed, but also to subjects' history of visual impairment, if any. Effective visualization strategies, for example, may be available to blindfolded and late blind subjects, but not to subjects who have been blind since birth (Poirier *et al.* 2007). It is also important to keep in mind findings concerning the crossmodal plasticity of the visually deafferentated brain. In particular, there is extensive evidence that occipital cortex is recruited for tactile and auditory processing when it is deprived of its standard sources of retinal input (Pascual-Leone & Hamilton 2001, Proulx *et al.* in press, Kupers *et al.* 2011, Kupers & Ptito 2014). This means that subjects who are blind or undergo prolonged visual deprivation may make rather different use of the information conveyed by stimulation in the substituting modality than normally sighted subjects. Contrary to possibility (b), but in keeping with (a), Kupers *et al.* (2006) found that, after one week of training with an electrotactile tongue display unit, transcranial magnetic stimulation (TMS) of occipital cortex evoked tactile sensations on the tongues of one late blind and three early blind subjects – 'short-lasting experiences of distinct tingling, varying in intensity, extent, and topography depending on the locus of the occipital cortex that was stimulated' (13257) – but only phosphenes in blindfolded subjects.<sup>1</sup>

---

<sup>1</sup> This result is consistent with the finding that TMS applied to occipital cortex induces experiences of touch that are referred to the fingertips in blind Braille readers (Ptito *et al.* 2008a). See Ortiz *et al.* 2011, however, for evidence that tactile stimulation elicits phosphenes in some late blind subjects.

Whether or not subjects have previously enjoyed visual experience seems also likely to affect how they learn to detect and disambiguate the spatial information present in SSD-mediated sensory stimulation. Unlike early blind subjects, blindfolded and late blind subjects may be able to exploit various crossmodal correspondences that obtain between the substituting modality and vision, for example, the link between tactile and visual shape or between auditory frequency and elevation in the visual field (Evans & Treisman 2010, Spence 2011). Consistent with this, studies by Kim and Zatorre (2008) have found that, *prior* to any training, blindfolded subjects can perform image identification tasks using the vOICE at an above-chance level so long as they are given an explicit explanation of the image-to-sound conversion rules.

## 2. Action Movement in Sensory Substitution

In tactile-visual sensory substitution (TVSS), low-resolution, gray-scale images from a video camera are converted by a human machine interface (HMI), pixel by pixel, into vibrotactile stimulations on the skin of one's back. A main finding in early experiments conducted by Paul Bach-y-Rita and colleagues was that subjects learned to discriminate the spatial layout of the scene in front of them via TVSS only when they had *active control* over movement of the video camera (White *et al.* 1970, Bach-y-Rita 1972). Subjects who received visual input passively experienced only a changing pattern of tactile stimulation:

When asked to identify static forms with the camera fixed, subjects have a very difficult time; but when they are free to turn the camera to explore the figures, the discrimination is quickly established. With fixed camera, subjects report experiences in terms of feelings on their backs, but when they move the camera over the displays, they give reports in terms of externally localized objects in front of them (White *et al.* 1970: 25).

Motivated by this finding, proponents of the sensorimotor contingency theory of perception have maintained that passively stimulated subjects do not learn to localize objects in surrounding space by means of TVSS because they are unable to master the sensorimotor contingencies that govern use of the device. '[A]ctive movement', Susan Hurley and Alva Noë write, 'is required in order for the subject to acquire practical knowledge of the change from sensorimotor contingencies characteristic of touch to those characteristic of vision and the ability to exploit this change skillfully' (2003:145; also see Noë 2004, Kiverstein 2010, and O'Regan 2011).<sup>2</sup>

---

<sup>2</sup> Sensorimotor contingencies are regularities that govern way moving in relation to the surrounding environment gives rise to changes in proximal sensory stimulation. Approaching an object, e.g., causes its retinal image to undergo expansion, which, in turn, causes a corresponding change in the pattern of retinal stimulation. Moving away from the object has the opposite effect. In the modality of touch, squeezing a soft sponge produces a distinctive pattern of tactile stimulations in one's hand. Squeezing a clod of dry clay produces quite another.

Hurley and Noë make two main claims about TVSS. The first claim is an answer to the question concerning the end-product of learning in sensory substitution. They argue that TVSS is best understood as a case of cortical deference in which activity in a blind subject's somatosensory cortex takes its 'qualitative expression' from the character of the externally rerouted *visual input* that it receives. TVSS, in other words, enables blind subjects to perceive the spatial layout of the distal environment in a phenomenologically vision-like manner – possibility (b) above. The second claim is an answer to the question concerning the nature of the learning process in sensory substitution. Subjects learn to make efficient use of TVSS only by acquiring knowledge of the sensorimotor contingencies associated with the prosthetic modality. 'The distinctively visual character of TVSS-perception stems from the way perceivers can acquire and use practical knowledge of the common laws of sensorimotor contingency that vision and TVSS-perception share' (2002:145).

It seems clear that there is significant overlap between the sensorimotor contingencies that respectively characterize ordinary, non-prosthetic vision and TVSS. This is because the HMI systematically converts the 2D image produced by the video camera, pixel by pixel, into a corresponding 2D array of vibrotactile stimulation on the skin. In consequence, the effects of camera movements on the structure of the former are mirrored by changes in the structure of the latter. Approaching an object, for example, causes its image in the camera to loom, which, in turn, causes its size in the vibrotactile array (functionally, the prosthetic retina) to increase. Retreating from the object causes its image to shrink, which, in turn, causes its size in the vibrotactile array to decrease. Panning to the right, causes the object's image to shift to the left, which, in turn, causes a structurally similar change in the pattern of vibrotactile stimulation on the skin. And so on.

Psychological and neuroscientific investigations of sensory substitution undertaken in the last decade provide reasons to be skeptical about this account. According to Hurley and Noë, activity in somatosensory cortex takes its 'qualitative expression' in TVSS from the character of the externally rerouted visual input that it receives. In other words, TVSS-perception is a case of somatosensory cortical deference. The TMS studies by Kupers *et al.* (2006), mentioned above, as well as brain imaging studies of cortical activation during visual-to-tactile sensory substitution tasks (for reviews, see Kupers *et al.* 2011, Kupers & Ptito 2014) present a rather different picture. In particular, they suggest, first, that in blind subjects occipital cortex is recruited to process the somatosensory inputs produced by a visual-to-tactile SSD and, second, that the resulting experience is, in at least some respects, phenomenologically touch-like in character: 'Our studies suggest that the qualitative character of the subject's experience is not determined by the area of cortex that is active (cortical dominance), but by the source of input to it (cortical deference)' (Kupers & Ptito 2014: 44). As most researchers would readily acknowledge, however, there are also salient contrasts between the representational content of TVSS-perception and the representational content of ordinary touch. Proficient users of TVSS, in particular, acquire the ability to discern the geometrical properties and locations of objects at a distance from the body in space. In this respect, the

representational content of TVSS-perception is akin to that of vision in normally sighted subjects.<sup>3</sup>

### 3. Distal Attribution and Causal Inference

There are also reasons to be skeptical that passively stimulated subjects fail to discriminate 3D spatial layout via TVSS because they lack the opportunity to learn the sensorimotor contingencies that govern use of the device. An alternative explanation is that subjects who do not control the camera's movement – and who are not otherwise attuned to its current position – are simply unable detect and make effective use of the *spatial information* present in the vibrotactile stimulations that they receive. In consequence they do not engage in 'distal attribution' (Loomis 1982, Epstein 1986): they do not attribute the cause of those stimulations to a three-dimensional scene in the external world.

Evidence that subjects with active control over camera movement do engage in distal attribution comes from the earliest experiments on TVSS by Bach-y-Rita and colleagues. Trained subjects, e.g., exhibited defensive, startle responses when objects loomed in the tactile array (Bach-y-Rita 1972: 98-99) and were spontaneously able to make sense of kinetic depth displays:

A modified version of the Metzger apparatus was used, consisting of a turntable on which two vertical white rods were mounted. This was rotated slowly before the camera and the subjects were asked to describe what they 'saw.' Some sighted subjects, upon first tactile presentation of this moving display, have spontaneously described it as moving in depth. Several blind subjects were given experience with a yoked pair of turntables. On one of these, an object was placed within view of the camera, while the subject turned the other freely, experiencing the transformations that the object underwent with the rotation. After an hour's experience with this equipment, they could report accurately the eccentric placement of two and three objects on the turntable, and could also experience rotation in depth with the Metzger display (White *et al.* 1970: 25).

In keeping with a recently influential Bayesian approach to perception in cognitive science (Mamassian *et al.* 2002, Kerstein & Yuille 2003, Shams & Beierholm 2010, Clark 2013, Hohwy 2013, Rescorla forthcoming), it is helpful, I would suggest, to think of distal attribution as the solution to a *causal inference* problem faced by the subject's perceptual systems. Patterns of proximal sensory stimulation significantly underdetermine their causal antecedents in the environment (hence, the so-called 'inverse optics' problem for vision). The central challenge faced by any perceptual system, on Bayesian models of perception, is to infer the most probable cause of a given pattern of proximal sensory stimulation on the basis of (1) the low-level cues

---

<sup>3</sup> I'm grateful to Julian Kiverstein for discussion of the TMS studies conducted by Kupers *et al.* 2006.

present in the pattern itself, i.e., properties of the pattern that are predictive of properties in the environment, and (2) pre-wired or learned assumptions about the environment’s statistical structure (the perceptual system’s ‘prior knowledge’ about the world).<sup>4</sup> The cues constrain the inference process from the bottom-up, while the assumptions constrain the process by, among other things, determining which way of integrating cues, in context, is statistically optimal (Knill 2007). The content of the perceptual state formed in response to a particular pattern of stimulation – the brain’s operative ‘hypothesis’ about the structure of the impinging environment – is the cause to which the highest probability is assigned given all the available endogenous and exogenous evidence. In the case of vision, this will normally be one of indefinitely many possible three-dimensional scenes.<sup>5</sup> The default hypothesis space for causal inference in everyday vision (the space of world states over which the posterior distribution is computed) is a *distant scene space*, in which different hypotheses correspond to different possible arrays of objects at a distance from the perceiver’s eyes. (One such hypothesis picks out the very scene present in front of the reader now.) By contrast, the default hypothesis space for causal inference in everyday touch is a *contacting object space*, in which different hypotheses correspond to different possible objects in contact with the surface of the perceiver’s body. Distal attribution in visual-to-tactile sensory substitution can be predicted to occur, when, from the standpoint of the perceiver’s perceptual systems, the most likely environmental cause of incoming vibrotactile or electrotactile stimulation – contrary to the default haptic interpretation – is a distant scene rather than an object of some sort touching the surface of her body.

The question to ask now is: What role does active control of camera movement play in enabling such radical re-shaping of the hypothesis space for touch in TVSS?

It is easiest to explain why such re-shaping is unlikely to occur when the subject is passively stimulated. When a subject outfitted with a TVSS device lacks control of the camera’s movements and is not otherwise able to register its position, all of the available endogenous and exogenous evidence is consistent with the default haptic interpretation of incoming sensory stimulation: the most likely cause of that stimulation is direct, bodily contact with an object of some kind. From a causal inference perspective, it is thus unsurprising that she does not learn to perceive

---

<sup>4</sup> This is a simplification. Bayesian models often incorporate other kinds of prior world knowledge. For discussion, see Geisler 2008. Two further points are important. First, perceptual inference is a non-conscious, subpersonal process. It is the perceiver’s brain rather than the perceiver herself that is confronting the causal inference problem described here. Second, Bayesian models do not assume that perceptual systems explicitly represent either the norms of Bayesian decision making or the various forms of prior knowledge imputed to them for purposes of explaining the formation of perceptual mental states. Bayesian models only assume that perceptual processes typically proceed in *accordance* with the principles of Bayesian decision making (see Burge 2010: 95-97 and Rescorla forthcoming).

<sup>5</sup> Although multistable perceptual experiences can occur in which the selected hypothesis alternates from one moment to the next, depending on the allocation of attention and other factors. Examples include the flip in depth assignments when viewing a drawing of a Necker cube or the reversal in perceived direction of rotation in the silhouette illusion.

spatially remote objects and features by means of TVSS. There is simply no reason for her perceptual systems to shift away from the hypothesis that the tactile stimulations she is receiving have an ordinary tactile cause. In consequence, although multiple sources of information about the visible scene are *present* in those stimulations – these derive from monocular cues in the 2D image such as height in the visual field, relative size, familiar size, linear perspective, motion parallax, shape-from-motion, and looming (White *et al.* 1970, Bach-y-Rita 1972) – the subject’s perceptual systems are unable to detect and exploit them.

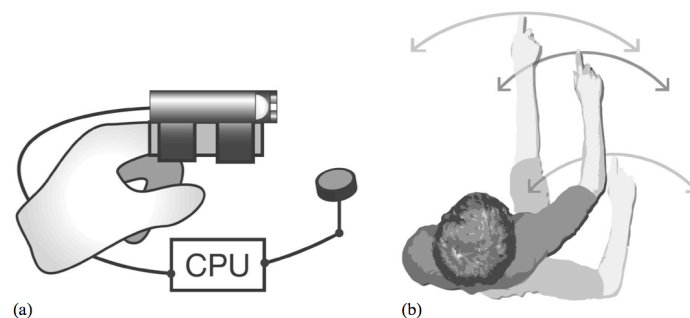
By contrast, when the subject has the ability to guide and keep track of the camera’s movement, she also has a significant amount of voluntary control over whether and how the vibrotactile stimulation she experiences undergoes change. In consequence, the situation is now one that conflicts with the default haptic interpretation: it is not typically possible to modify tactile stimulation on the surface of one’s back by moving a camera mounted on a tripod or on one’s head! The observed coupling between camera movement and vibrotactile stimulation is evidence that the latter’s cause resides outside of the hypothesis space for everyday touch.

In addition to evidence against the default haptic interpretation, there is now also additional evidence required for inference to a spatial layout in the hypothesis space of vision (or rather what constitutes the hypothesis space of vision in normally sighted subjects). To begin with, there are the aforementioned spatial cues present in the pattern of vibrotactile stimulation produced by the HMI. In addition, because the subject has control over how the camera moves she now also has access to real-time information about its body-relative position. This is important because the spatial cues present in the vibrotactile stimulation the subject is receiving are sources of information about the way objects fill out three-dimensional space *in front of the camera*. Height in the visual field, relative size, motion parallax, looming, etc. are all variations in image structure that are predictive of the distances of objects that reflect light to the camera’s lens. The HMI in TVSS works by converting these variations in image structure into corresponding variations in the structure of the vibrotactile array. To engage in causal inference to the way objects fill out space in front of her own body – to make use of the cues contained in the stimulation that she receives from the HMI as sources of *egocentric* spatial information – the subject thus needs to be able to keep track of the camera’s body-relative position, where it is located, for example, relative to her head, or torso, or hand. As Bach-y-Rita writes, ‘In the absence of motor control over the orientation of the sensory input, a person may have no idea from where the information is coming, and thus no ability to locate [its source] in space’ (2004: 90). Another key reason, then, that active control over the movements of the camera matters is that it generates information about the camera’s body-relative position necessary for causal inference to the way objects are arrayed in space at distance from the subject’s body. Such causal inference, on the present interpretation, is the essence of distal attribution.

I should emphasize that it is the availability of up-to-date and accurate information about how the camera is positioned in relation to the body that is important here, rather than the particular *way* in which the information is generated. In the case of

the subject who actively controls the camera's movements, the information is presumably based on proprioceptive signals and efference copy of motor commands to move the head or hand (depending, respectively, on whether the camera is mounted on a pair of eyeglasses or, as in the earliest experiments, on a tripod). But reliable information about body-relative camera position could, in principle, be made available in other ways to passively stimulated subjects.<sup>6</sup>

Siegle & Warren 2010 offer a similar interpretation of the role of self-produced arm movements in learning to make distance judgments using a simple visual-to-tactile SSD (Figure 1). The device in their experiments consisted of a single photodiode mounted on the index finger linked to a vibrating motor worn on the back. The motor was active when a subject pointed the photodiode in the direction of a light source and inactive otherwise. Subjects outfitted with the device were trained to sweep their arm back and forth in order to estimate the egocentric distance of a target light, which could be stationed at various locations along a 193-cm track in front of them. On Siegle and Warren's preferred Gibsonian interpretation, action is necessary for subjects to perform this task successfully not because it enables them to learn the sensorimotor contingencies associated with use of the SSD, but rather 'to reveal invariant information about the distal layout and to dissociate it from varying stimulation that depends on self-movement' (221). According to this 'invariance hypothesis', action generates proprioceptive information about pointing direction (generated by changes in the orientation of the torso and the joint angles of the shoulder and elbow), which when combined with the motor signal, is sufficient to triangulate the body-relative position of the distal target. Such triangulation, whether undertaken explicitly or implicitly, is clearly also available to subjects who control camera movement when using more sophisticated visual-to-tactile SSDs. They can use information about the camera's changing, body-relative orientation together with concurrent tactile stimulations to confirm that an object is 'out there' at a certain egocentrically defined distance and direction in space.



**Fig. 1** (a) The sensory substitution device used by Siegel & Warren 2010. Photodiode on the index finger responds to light from the target, driving a vibrating motor on the subject's back. (b) Pattern of arm movements participants were instructed to make when exploring the target. Reproduced from Siegel & Warren 2010.

<sup>6</sup> I am grateful to Fiona Macpherson and David Pence for discussion of this point.



The interpretation of the role of active camera control in learning to make use of TVSS offered above is a version of the invariance hypothesis. A main reason that subjects who control the camera's movements engage in distal attribution, I have argued, is that such control generates reliable information about the camera's body-relative position needed to exploit the spatial cues in incoming vibrotactile stimulations for purposes of egocentric object localization. Subjects do not need to acquire knowledge of the relevant sensorimotor contingencies.

#### **4. Bodily Action and Prism Adaptation**

The discussion in the last section suggests that we cannot straightforwardly infer from the apparent dependence of a given form of perceptual learning on self-produced bodily movement to the dependence of that form of learning on knowledge of sensorimotor contingencies, i.e., knowledge of the sensory consequences of action. Otherwise, we may overlook the Gibsonian possibility that it is not self-produced movement *per se* that makes the crucial difference to the learning process, but rather the perceptual or proprioceptive information to which the former gives rise.<sup>7</sup>

A final example illustrating this point comes from studies of how subjects adapt to prisms that reverse, invert, or laterally displace the retinal image (for overviews, see Rock 1966, Welch 1978). Richard Held and Alan Hein conducted an influential series of experiments in which participants wore laterally displacing prisms during either active or passive movement conditions (Held & Hein 1958, Hein & Held 1962, Held 1965). In the active movement condition, the subject moved her visible hand back and forth along a fixed arc in synchrony with a metronome. In the passive movement condition, the subject's hand was passively moved at the same rate by the experimenters. Although the overall pattern of visual stimulation was identical in both conditions, adaptation was reported only when subject's engaged in self-produced movement. Held and Hein used these findings to defend what in the literature has come to be known as the 'reafference theory' of adaptation. According to the reafference theory, subjects exhibit stable adapt to optical rearrangement only when they receive visual feedback from self-produced bodily movement, i.e., reafferent visual stimulation.

Contrary to reafference theory, subsequent experiments in the 1960s found that adaptation to lateral displacement is not restricted to situations in which subjects engage in movements that generate reafferent visual feedback, but can also take place when subjects receive visual feedback generated by passive effector or whole-body movements (Singer & Day 1966, Templeton *et al.* 1966, Fishkin 1969). Evidence was even garnered that prism adaptation is possible in the complete absence of motor action (Howard *et al.* 1965, Kravitz & Wallach 1969). In general, the extent to which adaptation occurs seems to depend not on the availability of reafferent stimulation from motor actions, as Held proposed, but rather on the

---

<sup>7</sup> For useful discussions of this point, see the commentaries on Gyr *et al.* 1979 and Campos 2000.

presence of either of two related kinds of information concerning ‘the presence and nature of the optical rearrangement’ (Welch 1978: 24). Following Welch, I shall refer to this alternative to the reafference theory as the *information hypothesis*.

One source of such information concerns the veridical directions of objects from the observer (Rock 1966; chapters 2-4). Normally, when engaging in forward locomotion, the apparent radial direction of an object straight ahead of the body remains constant, while the apparent radial directions of objects to either side undergo change. This pattern also obtains when the observer wears prisms that displace the retinal image to side. Hence, as Rock writes, ‘an object seen through prisms which retains the same radial direction as we approach must be seen to be moving in toward the sagittal plane’ (1966: 105). On Rock’s view, at least some forms of prism adaptation can be explained by our ability to detect and exploit such stable sources of spatial information in locomotion-generated patterns of optic flow.

Another more effective source of information for prism adaptation is the registered discrepancy between seen and proprioceptively experienced limb position (Wallach 1968). Proponents of the information hypothesis have found that, when this conflict is made conspicuous, passively moved (Melamed *et al.* 1973), involuntarily moved (Mather & Lackner 1975), and even immobile subjects (Kravitz & Wallach 1966) exhibit significant adaption. Although active bodily movement is unnecessary for adaptation to occur, it provides subjects with especially salient and precise information about the discrepancy between sight and touch (Moulden 1971): subjects are able proprioceptively to determine the location of a moving limb much more accurately than a stationary or passively moved limb.

## 5. Conclusion

Sources of evidence for the information hypothesis discussed above indicate that adaptation to displacing prisms can occur in the absence of self-produced movement. The information hypothesis, however, predicts that action will *facilitate* adaptation when it generates information either about the world (in particular, the real radial directions of objects from the perceiver) or about the perceiver’s body (in particular, the real positions of her limbs) that conflicts with the way things look as a result of prismatic displacement. ‘According to this view’, Rock writes, ‘[active] movement is important only because it allows for certain kinds of information to be registered, not because movement *per se* is necessary’ (1966: 42).

In this chapter, my aim has been to show that something analogous holds true of the learning process in tactile-visual sensory substitution (TVSS). In particular, I have argued that active control over camera movement facilitates distal attribution in TVSS not because it enables subjects to master the laws of sensorimotor contingency governing use of the device, but rather because it generates proprioceptive and efference-copy based information about the camera’s changing body-relative position. Such information is important to the emergence of distal attribution for two main reasons. First, without knowledge that the vibrotactile stimulations she is

receiving can be modified by camera movement, the subject’s perceptual systems have no reason to budge from the default haptic interpretation of those stimulations. The coupling between changes in camera position and changes in sensory stimulation is evidence that the cause of the latter resides outside the ordinary hypothesis space for touch. Second, having the ability to keep track of the camera’s body-relative position enables the subject to exploit the spatial cues present in vibrotactile stimulation for causal inference to the egocentrically specified distances and directions of objects in space around her. Such causal inference, on the interpretation I have defended here, is the basis of distal attribution.<sup>8</sup>

## References

- Amedi A., Stern, W., Camprodon, J., Bermpohl, F., Merabet, L., & Rotman, S. (2007), ‘Shape Conveyed by Visual-to-Auditory Sensory Substitution Activates the Lateral Occipital Complex’, *Nature Neuroscience*, 10: 687–689.
- Bach-y-Rita, P. (1972), *Brain Mechanisms in Sensory Substitution* (New York, Academic Press).
- Bach-y-Rita, P. (2004), ‘Tactile Sensory Substitution Studies’, *Annals of the New York Academy of Sciences*, 1013: 83-91.
- Burge, T. (2010), *Origins of Objectivity* (Oxford, Oxford University Press).
- Campos, J., Anderson, D., Barbu-Roth, M., Hubbard, E., Hertenstein, M., & Witherington, D. (2000), ‘Travel Broadens the Mind’, *Infancy*, 1: 149-219.
- Capelle, C., Trullemans, C., Arno, P., & Veraart, C. (1998), ‘A Real-Time Experimental Prototype for Enhancement of Vision Rehabilitation Using Auditory Substitution’, *IEEE Transaction on Biomedical Engineering*, 45: 1279-1293.
- Chebat, D., Schneider, F., Kupers, R., & Püto, M. (2011), ‘Navigation with a Sensory Substitution Device in Congenitally Blind Individuals’, *Neuroreport*, 22: 342-47.
- Clark, A. (2013), ‘Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science’, *Behavioral and Brain Sciences*, 36: 181-253.
- Evans, K. & Treisman, A. (2010), ‘Natural Cross-Modal Mappings Between Visual and Auditory Features’, *Journal of Vision*, 10: 1-12.
- Fishkin, S. (1969), ‘Passive vs. Active Exposure and Other Variables Related to the Occurrence of Hand Adaptation to Lateral Displacement’, *Perceptual and Motor Skills*, 29: 291-297.
- Geisler, W. (2008), ‘Visual Perception and the Statistical Properties of Natural Scenes’, *Annual Review of Psychology*, 59: 167-192.
- Gyr, J., Siley, R., & Henry, A. (1979), ‘Motor-Sensory Feedback and the Geometry of Visual Space’, *Behavioral and Brain Sciences*, 2: 59-94.
- Held, R. & Hein, A. (1958), ‘Adaptation of Disarranged Hand-Eye Coordination Contingent upon Re-Afferent Stimulation’, *Perceptual and Motor Skills*, 8: 87-90.
- Hein, A. & Held, R. (1962), ‘A Neural Model for Labile Sensorimotor Coordinations’, in E. Bernard & M. Kare (eds), *Biological Prototypes and Synthetic Systems*, volume 1 (New York, Plenum Press), 71-74.
- Held, R. (1965), ‘Plasticity in Sensory-Motor Systems’, *Scientific American*, 213: 84-94.
- Hohwy, J. (2013), *The Predictive Brain* (Oxford, Oxford University Press).

---

<sup>8</sup> For helpful discussions, I am grateful to Amir Amedi, Derek Brown, Jonathan Cohen, Joshua Downing, Julian Kiverstein, Fiona Macpherson, Mohan Matthen, Paul Noordhof, Kevin O’Regan, David Pence, Michael Proulx, and Jamie Ward.

- Howard, I., Craske, B., & Templeton, W. (1965), 'Visuomotor Adaptation to Discordant Exafferent Stimulation', *Journal of Experimental Psychology*, 70: 189-191.
- Hurley, S. & Noë, A. (2003), 'Neural Plasticity and Consciousness', *Biology and Philosophy*, 18; 131-168.
- Kersten, D. & Yuille, A. (2003), 'Bayesian Models of Object Perception', *Current Opinion in Neurobiology*, 13: 150-158.
- Kim, J. & Zatorre, R. (2008), 'Generalized Learning of Visual-to-Auditory Substitution in Sighted Individuals', *Brain Research*, 1242: 263-275.
- Kiverstein, J. (2010), 'Sensorimotor Knowledge and the Contents of Experience', in N. Gangopadhyay, M. Madary, & F. Spicer (eds), *Perception, Action, and Consciousness* (Oxford, Oxford University Press), 257-274.
- Knill, D. (2007), 'Learning Bayesian Priors for Depth Perception', *Journal of Vision*, 7: 1-20.
- Kravitz, J. & Wallach, H. (1966), 'Adaptation to Displaced Vision Contingent upon Vibrating Stimulation', *Psychonomic Science*, 6: 465-466.
- Kupers, R., Fumal, A., de Noordhout, A., Gjedde, A., Schoenen, J., & Püto, M. (2006), 'Transcranial Magnetic Stimulation of the Visual Cortex Induces Somatotopically Organized Qualia in Blind Subjects', *Proceedings of the National Academy of Sciences*, 103: 13256-13260.
- Kupers, R., Pietrini, P., Ricciardi, E., & Püto, M. (2011), 'The Nature of Consciousness in the Visually Deprived Brain', *Frontiers in Psychology*, 2.
- Loomis, J. (1982), 'Distal Attribution and Presence', *Presence*, 1: 113-119.
- Mamassian, P., Landy, M., & Maloney, L. (2002), 'Bayesian Modeling of Visual Perception', in R. Rao, B., Olshausen, & M. Lewicki (eds), *Probabilistic Models of the Brain* (Cambridge, MA, MIT Press).
- Mather, J. & Lackner, J. (1975), 'Adaptation to Visual Rearrangement Elicited by Tonic Vibration Reflexes', *Experimental Brain Research*, 24: 103-105.
- Meijer, P. B. L. (1992), 'An Experimental System for Auditory Image Representations', *IEEE Transactions on Biomedical Engineering*, 39: 112-121.
- Melamed, L., Halay, M., & Gildow, J. (1973), 'Effect of External Target Presence on Visual Adaptation with Active and Passive Movement', *Journal of Experimental Psychology*, 98: 125-130.
- Merabet, L. & Pascual-Leone, A. (2009), 'Neural Reorganization Following Sensory Loss: The Opportunity of Change', *Nature Reviews Neuroscience*, 11: 44-52.
- Moulden, B. (1971), 'Adaptation to Displaced Vision: Reafference is a Special Case of the Cue-Discrepancy Hypothesis', *The Quarterly Journal of Experimental Psychology*, 23: 113-117.
- Noë, A. (2004), *Action in Perception* (Cambridge, MA, MIT Press).
- O'Regan, J. (2011), *Why Red Doesn't Sound Like a Bell: Explaining the Feel of Consciousness* (Oxford, Oxford University Press).
- O'Regan, J., & Noë, A. (2001), 'A Sensorimotor Account of Vision and Visual Consciousness', *Behavioral and Brain Sciences*, 24: 939-973.
- Ortiz, T., Poch, J., Santos, J., Requena, C., Martínez, A., *et al.* (2011), 'Recruitment of Occipital Cortex During Sensory Substitution Training Linked to Subjective Experience of Seeing in People with Blindness', *PLoS ONE* 6: e23264.
- Pascual-Leone, A. & Hamilton, R. (2001), 'The Metamodal Organization of the Brain', *Progress in Brain Research*, 134: 427-445.
- Poirier, C., De Volder, A., & Scheiber, C. (2007), 'What Neuroimaging Tells Us About Sensory Substitution', *Neuroscience and Biobehavioral Reviews*, 31: 1064-70.
- Proulx, M., Brown, D., Pasqualotto, A., & Meijer, P. (in press), 'Multisensory Perceptual Learning and Sensory Substitution', *Neuroscience and Biobehavioral Reviews*.

- Ptito, M., Moesgaard, S., Gjedde, A., & Kupers, R. (2005), 'Cross-Modal Plasticity Revealed by Electrotactile Stimulation of the Tongue in the Congenitally Blind', *Brain*, 128: 606–614.
- Ptito, M., Fumal, A., de Noordhout, A. M., Schoenen, J., Gjedde, A., & Kupers, R. (2008a), 'TMS of the Occipital Cortex Induces Tactile Sensations in the Fingers of Blind Braille Readers', *Experimental Brain Research*, 184: 193-200.
- Ptito, M., Schneider, F., Paulson, O., & Kupers, R. (2008b), 'Alterations of the Visual Pathways in Congenital Blindness', *Experimental Brain Research*, 187: 41-49.
- Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C., De Volder, A. (2005), 'Cross-modal Activation of Visual Cortex During Depth Perception Using Auditory Substitution of Vision', *Neuroimage*, 26: 573-80.
- Rescorla, M. (forthcoming), 'Bayesian Perceptual Psychology. To appear in M. Matthen (ed.), *The Oxford Handbook of the Philosophy of Perception* (Oxford, Oxford University Press).
- Rock, I. (1966), *The Nature of Perceptual Adaptation* (New York, Basic Books).
- Sampaio, E., Maris, S., Bach-y-Rita P. (2001), 'Brain Plasticity: "Visual" Acuity of Blind Persons Via the Tongue,' *Brain Research*, 908: 204–207.
- Shams, L. & Beierholm, U. (2010), 'Causal Inference in Perception', *Trends in Cognitive Sciences*, 14: 425-432.
- Siegle, J. & Warren, W. (2010), 'Distal Attribution and Distance Perception in Sensory Substitution', *Perception*, 39: 208-223.
- Singer, G. & Day, R. (1966), 'Spatial Adaptation and Aftereffect with Optically Transformed Vision', *Journal of Experimental Psychology*, 71: 725-731.
- Spence, C. (2011), 'Crossmodal Correspondences: A Tutorial Review', *Attention, Perception, and Psychophysics*, 73: 971-995.
- Templeton, W., Howard, I., & Lowman, A. (1966), 'Passively Generated Adaptation to Prismatic Distortion', *Perceptual and Motor Skills*, 22: 140-142.
- Wallach, H. (1968), 'Informational Discrepancy as a Basis of Perceptual Adaptation', in S. Friedman (e.), *The Neuropsychology of Spatially Oriented Behaviour* (Homewood, Illinois, Dorsey Press), 209-230.
- Welch, R. (1978), *Perceptual Modification: Adapting to Altered Sensory Environments* (New York, Academic Press).
- White, B., Saunders, F., Scadden, L., Bach-y-Rita, P., & Collins, C. (1970), 'Seeing with the Skin', *Perception and Psychophysics*, 7: 23–27.